Title	KOLMOGOROV COMPLEXITY AND P-PRINTABLE SETS(Metamathematics and it's applications)
Author(s)	今田, 宏司; 篠田, 寿一
Citation	数理解析研究所講究録 (1995), 930: 112-119
Issue Date	1995-11
URL	http://hdl.handle.net/2433/59942
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

数理解析研究所講究録
930巻 1995年 112-119

KOLMOGOROV COMPLEXITY AND P-PRINTABLE SETS

今田宏司 (KOJI IMADA), 篠田寿一 (JUICHI SHINODA)
名古屋大学・人間情報学研究科

1. Generalized Kolmogorov complexity

まず, $\Sigma = \{0,1\}$ とし, Σ 上の有限列の全体を Σ^* とする. $x \in \Sigma^*$ に対し, その長さを $|x|$ で表す. その他の P, NP 等の Notation については Complexity Theory で用いられる通常のものとする.

Kolmogorov complexity とは, 対象の本質的な情報量を捉えようとするものであり, Kolmogorov, Chaitin, Solomonoff 等によって独立に発見された概念である. finite string の Kolmogorov complexity はその finite string を出力する最も短い program の長さとして定義される. 直感的に, ある string がそれ自身よりも短い program によって生成されるならばその string は冗長な情報を持っている (i.e. 圧縮する事ができる) ことがわかる. そして圧縮する (i.e. *compress*) 事が出来ないならば, その string は *random* であるという.

Kolmogorov complexity を randomness の定義として用いることの限界は, 元の string をその最も短い program が生成するときの computation に使われる資源 (time, space) を全く制限しないことである. そこで, Kolmogorov complexity の time-bounded version が Ko[Ko86], Sipser[Sip83], Hartmanis[Har88] らによって考えられた.

Hartmanis は二つのパラメーターを持つ Kolmogorov complexity(*generalized Kolmogorov complexity* と呼ぶ) を提唱した. これは string がどれだけ圧縮されるかだけでなく, 如何に早く展開 (i.e. *restore*) できるかということも考慮する.

ここでは Hartmanis の以下の定義を用いる.

定義 1.1. M_u を Turing Machine とし, $g, G : \boldsymbol{N} \to \boldsymbol{N}$ とする. このとき

$$K_u[g(n), G(n)] = \left\{ x \in \Sigma^* \;\middle|\; (\exists y \in \Sigma^*) \begin{bmatrix} M_u(y) \text{ halts within } G(|x|)\text{-steps} \\ M_u(y) = x \;\; \& \;\; |y| \leq g(|x|) \end{bmatrix} \right\}$$

y を *compressed string*, x を *restored string* と呼ぶ. そして $g(n)$ を *compression*, $G(n)$ を *restoration time* と呼ぶ.

定理 1.1 (Hartmanis). 万能 Turing machine M_u を適当にとって, すべての Turing machine M_v に対し

$$K_v[g(n), G(n)] \subseteq K_u[g(n) + c, c\, G(n) \log G(n) + c]$$

が成り立つようにできる. ここに $c = c_v$ は M_v から定まる定数である.

以後, 定理のような universal Turing machine M_u を固定し, 添字 u を省略して

$$K[g(n), G(n)] = K_u[g(n), G(n)]$$

と書く.

定義 1.2. $S \subseteq \Sigma^*$ が small generalized Kolmogorov complexity をもつとは

$$S \subseteq K[k \log n, n^k] \quad \text{for some } k$$

であることをいう.

これらは, 任意の oracle A に相対化できる. いま

$$K^A[g(n), G(n)] = K_u^A[g(n), G(n)]$$

とする. ここで M_u^A は定理 1.1 を A に相対化して得られる A-universal Turing machine である.

定義 1.3. (1) S が small generalized Kolmogorov complexity relative to A をもつとは

$$S \subseteq K^A[k \log n, n^k], \quad \text{for some } k$$

であることをいう.

(2) S が small generalized Kolmogorov complexity relative to itself をもつとは

$$S \subseteq K^S[k \log n, n^k], \quad \text{for some } k$$

であることをいう.

2. P-PRINTABLE SETS AND GENERALIZED KOLMOGOROV COMPLEXITY

定義 2.1. S が sparse であるとは，適当に多項式 $p(n)$ をとれば

$$|\{x \in S : |x| = n\}| \leq p(n)$$

となることをいう.

そして sparse set と generalized Kolmogorov complexity については次のような関係がある.

補題 2.1. S が small generalized Kolmogorov complexity (relative to A) をもてば，S は sparse である.

証明. $x \in S$ かつ $|x| = n$ ならば，定義より

$$|y| \leq k \log n, \quad M_u(y) = x$$

となる $y \in \Sigma^*$ が存在する．長さが $k \log n$ 以下の Σ^* の元の総数は

$$1 + 2 + 2^2 + \cdots + 2^{k \log n} = 2^{k \log n + 1} - 1 \leq 2n^k - 1$$

であるから，$p(n) = 2n^k - 1$ とおけば，

$$|\{x \in S : |x| = n\}| \leq p(n)$$

である. □

sparse set は，すべての n にたいしてサイズが，n の polynomial のテーブルに，長さ n 以下の string を格納することが出来るという有用な性質を持っている．S を sparse set とし，十分大きな n に対し，サイズが n の polynomial のテーブルに長さ n 以下の S の string をすべて格納することで，S 自身の complexity がいかに大きくても必要とされる S のすべての情報に迅速にアクセスすることが出来る.

しかし，そのようなテーブルを作る事の complexity は S を認識することの complexity よりもはるかに高いと思われる．そのような S のテーブルを作ることの complexity が，S 自身の complexity よりもあまり高くないとき，S を *self P-printable* であるという．そして S のテーブルを作ることが難しくないならば，*P-printable* であると言う．すなわち

定義 2.2. S が P-printable であるとは，適当な PTIME bounded Turing machine M が存在して，すべての n に対して

$$M(0^n) = x_1 x_2 \cdots x_L, \quad \text{ただし } \{x \in S : |x| = n\} = \{x_1, x_2, \ldots x_L\}$$

となることをいう．すなわち M に 0^n を入力したとき，n の多項式時間内に長さ n の S の元をリストにして出力することをいう．

定義 2.3. S が P^A-printable であるとは，適当な PTIME bounded oracle Turing machine M^X が存在して，すべての n に対して

$$M^A(0^n) = x_1 x_2 \cdots x_L, \quad \text{ただし } \{x \in S : |x| = n\} = \{x_1, x_2, \ldots x_L\}$$

となることをいう．S が P^S-printable であるとき self P-printable であるという．

P-printable set はあきらかに sparse であってしかも P に含まれる．

P-printable set と generalized Kolmogorov complexity については次のような定理 [AR88] がある．

定理 2.1 ([AR88]). 以下は同値．

(1) S は P-printable
(2) S は sparse かつ P-time computable な ranking function を持つ．
(3) S はある tally set in P に P-isomorphic.
(4) $S \subseteq K[k \log n, n^k]$ かつ $S \in P$

ここで ranking function とは次のように定義されるものである．

定義 2.4. 任意の集合 $L \subseteq \Sigma^*$ に対し，L の ranking function $r_L : \Sigma^* \to \boldsymbol{N}$ とは次の条件を満たす物である．

$$r_L(x) = |\{w \in L : w \leq x\}|$$

命題 2.1. 任意の Turing machine M_v, k に対し，

$$K_v[k \log n, n^k] \in P$$

上の定理と命題を組み合わせることで次の Corollary が導かれる．

系 2.1. 以下は同値

(1) $A \subseteq K[k \log n, n^k]$ となる k が存在する．(i.e. A は small generalized Kolmogorov complexity を持つ)

(2) A はある tally set に P-isomorphic

定理 2.1 の (1) と (4) の同値性は[BB86], [HH86] でも証明されているし，以下のように直接証明することができる．

補題 2.2. S が P-printable であるための必要十分条件は，次の (i), (ii) が成り立つことである．

(i) $S \in \mathrm{P}$,

(ii) S は small generalized Kolmogorov complexity をもつ．

証明. S を P-printable set とすると，適当な PTIME bounded Turing machine M をとれば，$M(0^n)$ が $\{x \in S : |x| = n\}$ の元をリストアップしたものになっている．いま $M(0^n)$ の計算が n^k steps 以内に終了するとする．まず $S \in \mathrm{P}$ であることを示す．$x \in \Sigma^*$, $|x| = n$ とする．$M(0^n)$ を計算した結果

$$w = x_1 x_2 \cdots x_L$$

が出力されたとする．w の先頭から n ビットずつ見ていくことにより，$x = x_i$ となる i $(1 \leq i \leq L)$ が存在すれば $x \in S$ であり，そうでなければ $x \notin S$ である．したがって x が S に属するか否かは $n^k + n^t \cdot L \leq (1 + n^t)n^k$ steps 以内で決定できる．ここに n^t は x と x_i を比較するために要する時間である．

次に S が small generalized Kolmogorov complexity を持つことを示す．$|n|$ が与えられたとき，n steps で 0^n を生成することができる．これを M に入力し n^k steps で $\{x \in S : |x| = n\}$ の元のリスト

$$x_1 x_2 \cdots x_L$$

を得る．このリストの i 番目の要素 x_i を出力するために ni steps を要する．したがって $x \in S$, $|x| = n$ に対して，$y = \langle |n|, i \rangle$ とおくと，x は y から $n + n^k + n \cdot L \leq n^{k+2}$ steps 以内に計算できる．したがって

$$S \subseteq K_M[4 \log n, n^{k+2}]$$

となり，定理 1.1 により S は small generalized Kolmogorov complexity をもつ．

$S \in \mathrm{P}$ かつ S が small generalized Kolmogorov complexity をもつとすると，

$$S \subseteq K[k \log n, n^k]$$

となる k が存在する．$|y| \leq k \log n$ なる $y \in \Sigma^*$ に対して，$u(y)$ を n^k steps 計算し，そのときの出力を x とするとき，$x \in S$，$|x| = n$ ならば，出力テープに x を書き出す．すべての y ($|y| \leq k \log n$) についてこれを実行すれば，$\{x \in S : |x| = n\}$ の元をリストアップすることができる．$|\{y \in \Sigma^* : |y| \leq k \log n\}| = 2n^k - 1$ であることに注意すれば，これに要する計算時間は $(2n^k - 1) \cdot (n^k + p(n))$ である．ここで $p(n)$ は，$|x| = n$ なる $x \in \Sigma^*$ に対して，$x \in S$ か否かを判定するために要する時間である．したがって S は P-printable である．□

同様の証明により次が得られる．

補題 2.3. S が self P-printable であるための必要十分条件は，S が small generalized Kolomogorov complexity relative to itself をもつことである．

定理 2.1 より，small generalized Kolmogorov complexity を持つ集合はすべて sparse であることがわかる．けれどもすべての sparse set が small Kolmogorov complexity を持つわけではない．それどころか，すべての polynomial よりも大きい任意の time constructible function $T(n)$ に対し，$K[k \log n, n^k]$ の subset でない sparse set in DTIME($T(n)$) が存在することを示すのは難しくない．

更に任意の $S(n) = o(n)$, recursive function $T(n)$ に対し，$K[S(n), T(n)]$ の subset でない nonrecursive sparse set が存在する．このような集合の例としては すべての長さ n に対し，ただひとつの Kolmogorov-random string からなる集合を考えればよい．

ただし，P に属する sparse set がすべて small generalized Kolmogorov complexity を持つ (定理 2.1 より P-printable) か否かは未解決である．

3. P-PRINTABLE でない SPARSE SET が P の中に存在するか？

この問題に関しては次の定理が [AR88] で証明されている．

定理 3.1. 以下の条件は同値

(1) P-printable でない sparse set in P が存在する．
(2) P-printable でない sparse set in DLOG が存在する．
(3) FewP $-$ P に属する sparse set が存在する．

ここで FewP とは NP に属する集合で，accepting computation path の個数が入力の長さの多項式で押さえられるものの class である．当然 $\mathrm{P} \subseteq \mathrm{FewP} \subseteq \mathrm{NP}$ である．

[BGS75] により，$\mathrm{P}^A = \mathrm{NP}^A$ となる oracle A が構成されている．この A に対しては，上の定理 (を A に相対化したもの) から，P^A-printable でない sparce set は P^A の中には存在しないことが分かる．

一方，次の定理から，適当な oracle S をとって，P^S-printable でない sparse set が P^S の中に存在するようにもできる．

定理 3.2. sparse な recursive set で，self P-printable でないものが存在する．したがって，sparse であるが，small Kolmogorov complexity to itself をもたない集合が存在する．

証明. PTIME bounded oracle Turing machine 全体の recursive enumeration を $\{M_e^X, p_e\}_{e \in N}$ とする．ただし p_e は M_e^X の running time をおさえる多項式である．帰納法で増加列

$$l_0 < l_1 < \ldots < l_e < \ldots$$

と $S_e : \{x : |x| < l_e\} \to \{0,1\}$ を以下のように定義する．

Stage 0. $l_0 = 0$, $S_0 = \emptyset$ とする．

Stage $e+1$. l_e, S_e がすでに定義されていると仮定して，

$$l_e \leq n, \quad 2p_e(n) < 2^n$$

をみたす最小の n をとり $l_{e+1} = 2^n$ とする．$S_e * 0$ を

$$(S_e * 0)(x) = \begin{cases} S_e(x) & \text{if } |x| < l_e \\ 0 & \text{if } |x| \geq l_e \end{cases}$$

によって定め，$w = M_e^{S_e * 0}(0^n)$ とする．

$$w = x_1 x_2 \cdots x_L, \quad |x_1| = |x_2| = \cdots = |x_{L-1}| = n, \quad |x_L| \leq n$$

であるとき，$L \leq p_e(n)$ である．また $M_e^{S_e * 0}(0^n)$ の計算中に oracle $S_e * 0$ に query する回数は高々 $p_e(n)$ である．$2p_e(n) < 2^n$ であるから，長さ n の列で $x_1, x_2, \ldots, x_L$ のすべてと異なり $M_e^{S_e * 0}(0^n)$ の計算中に query されないものが存在する．その最

初のものを x_0 とする．そして

$$S_{e+1}(x) = \begin{cases} S_e(x) & \text{if } |x| < l_e \\ 1 & \text{if } x = x_0 \\ 0 & \text{if } l_e \leq |x| < l_{e+1} \ \ \& \ \ x \neq x_0 \end{cases}$$

によって S_{e+1} を定める．

このように $\{l_e\}_{e\in N}$, $\{S_e\}_{e\in N}$ を定め

$$S = \bigcup_{e\in N} S_e$$

とおけば，

$$M_e^S(0^n) = M_e^{S_e * 0}(0^n) = w = x_1 x_2 \cdots x_L, \quad x_0 \neq x_1, x_2 \ldots, x_L, \quad x_0 \in S$$

が成り立つ．したがって $M_e^S(0^n)$ は $\{x \in S : |x| = n\}$ の元をリストアップすることはない．よって S は self P-printable ではない．S が sparse であることは，各 n に対して長さ n の列は高々 1 個しか S に入れられないことから明らかである．□

上の証明を少し変更することにより，self P-printable でない sparse set が $2^{\aleph_0}$ 個存在することもいえる．一方 tally set は $2^{\aleph_0}$ 個存在するから，self P-printable set はやはり $2^{\aleph_0}$ 個存在する．

References

[AR88] Eric W. Allender and Roy S. Rubinstein. P-printable sets. In *SIAM Journal on Computing*, volume 17, pages 1193–1202. SIAM, 1988.

[BB86] J. Baclcazar and R. Book. On generalized kolmogorov complexity. In *Acta Inform.*, volume 23, pages 679–688. 1986.

[BGS75] Theodore Baker, John Gill, and Robert Solovay. Relativizations of the $\mathcal{P} =? NP$ question. In *SIAM Journal on Computing*, volume 4, pages 431–442. SIAM, 1975.

[Har88] J. Hartmanis. Generalized kolmogorov complexity and the structure of feasible computations. In *Proceedings of 25th IEEE Symposium on Foundations of Computer Science*, pages 439–445. IEEE, 1988.

[HH86] J. Hartmanis and L. Hemachandra. On sparse oracles separating feasible complexity classes. In *Proceedings of 3rd Annual Symposium on Theoretical Aspects of Computer Science*, pages 321–333. Springer-Verlag New York Inc., 1986.

[Ima95] Koji Imada. Polynomial isomorphism type and polynomial one-one reduction. Master's thesis, Nagoya University, Japan, 1995.

[Ko86] K. Ko. On the notion of infinite pseudorandom sequences. In *Theoretical Computer Science*, volume 48, pages 9–13. North-Holland, 1986.

[Kur85] Stuart A. Kurtz. Sparse sets in NP-P: Relativizations. In *SIAM Journal on Computing*, volume 14, pages 113–119. SIAM, 1985.

[Sip83] M. Sipser. A complexity theoretic approach to randomness. In *Proceedings of 17th Annual ACM Symposium on Theory of Computing*, pages 330–335. ACM, 1983.